AFRL-HE-BR-TR-2004-0188

**United States Air Force**
**Research Laboratory**

# METHODOLOGY FOR EVALUATING THE SIMULATOR FLIGHT PERFORMANCE OF PILOTS

Jennifer K. Smith
John A. Caldwell

HUMAN EFFECTIVENESS DIRECTORATE
BIOSCIENCES AND PROTECTION DIVISION
FATIGUE COUNTERMEASURES BRANCH
2485 GILLINGHAM DRIVE
BROOKS CITY-BASE TX 78235-5105

December 2004

## NOTICES



//SIGNED//
JENNIFER K. SMITH
Project Scientist

//SIGNED//
F. WESLEY BAUMGARDNER, Ph.D.
Deputy, Biosciences and Protection Division

| REPORT DOCUMENTATION PAGE | *Form Approved* **OMB No. 0704-0188** |
|---|---|



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| December 2004 | Interim | 01-04-2003 – 01-12-2003 |

**4. TITLE AND SUBTITLE**
Methodology for Evaluating the Simulator Flight Performance of Pilots

**5a. CONTRACT NUMBER**
NA

**5b. GRANT NUMBER**
NA

**5c. PROGRAM ELEMENT NUMBER**
62022F

**6. AUTHOR(S)**
Smith, Jennifer K. & Caldwell, John A., Jr.

**5d. PROJECT NUMBER**
7757

**5e. TASK NUMBER**
P9

**5f. WORK UNIT NUMBER**
04

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Biosciences and Protection Division
2485 Gillingham Dr.
Brooks City-Base, TX 78235-5105

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSORING/MONITORING AGENCY REPORT NUMBER**
AFRL-HE-BR-TR-2004-0188

**12. DISTRIBUTION AVAILABILITY STATEMENT**
Approved for public release: Distribution is unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
The type of research that investigates operational tasks such as flying an aircraft or flight simulator is extremely useful to the Air Force's operational community because the results apply directly to "real-world" settings. However, the task of collecting, processing, and analyzing flight data involves handling a tremendous amount of information; therefore, it is vital to develop a precise methodology in order to successfully evaluate measures of pilot performance. A defined methodology was developed and implemented in the evaluation of flight data collected from the Link F-117A Weapon System Training Simulator during a fatigue study at Holloman Air Force Base, New Mexico. The steps involved in processing the flight data, including data collection, reduction, and analysis, are described in this paper. This methodology enabled the first ever objective quantification of fatigue effects on Air Force fighter pilots. The procedure revealed that the fatigue from 20 or more hours of sleep loss significantly impacted objectively measured basic piloting skills, including the ability to maintain headings, altitudes, airspeeds, etc. Fatigue affected some maneuvers more than others. This flight scoring methodology can be applied to digitized flight data collected from both simulators and aircraft in order to objectively measure the effects of stressors and interventions on flight performance.

**15. SUBJECT TERMS**
Flight data, Air Force, flight maneuvers, F-117, fatigue, simulator, methodology, pilots, performance, RMSE

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | UNCLASSIFIED | 25 | 1Lt Jennifer K. Smith |
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | | | **19b. TELEPONE NUMBER** ***(Include area code)*** 210-536-4777 |

TABLE OF CONTENTS



LIST OF FIGURES



LIST OF TABLES

# INTRODUCTION

Evaluating pilot performance is a difficult and complicated task due to the dynamic nature of the aviation environment. Throughout history, various efforts have explored the best techniques and performance measurements to effectively measure flight performance. Subjective and objective evaluations have traditionally been used to evaluate performance (Rehmann, 1982). Subjective evaluations often receive criticism because of the potential for observer biases and variability amongst observers. Also, it is difficult to quantify subjective ratings, and the rating scales often are not detailed enough to accurately reflect the behavior being evaluated. However, subjective rating scales can be useful for tasks that are difficult to measure such as decision making and mood evaluations. This was evident in a study by Connolly, Blackwell, and Lester (1989) where raters assigned a subjective score to the pilot's decision making ability. The researchers found this a useful tool for evaluating the pilot's judgment—an aspect of aviation job performance that is not currently amenable to more objective assessment. In the case of aircraft control skill, most evaluations have centered on the use of objective measures because they are quantifiable and involve identifiable standards against which performance can be measured (Mixon and Moroney, 1981).

The United States Air Force focused on a variety of objective Automated Performance Measurement (APM) tools beginning in 1968 in an attempt to increase the objectivity and reliability of pilot performance measurement (Fuller, Waag, and Martin, 1980). The APM tools became of particular interest in training environments such as the Undergraduate Pilot Training (UPT) program where objective evaluation techniques were needed to measure the student pilot's performance (Knoop and Welde, 1973). The

United States Army also sought to find objective, quantitative methods to evaluate pilot performance in rotary wing aircraft (Billings, Eggspuehler, Gerke and Chase, 1968). Although initial efforts to precisely quantify piloting skill were less than optimal, better methods for collecting, measuring, and analyzing objective flight performance data evolved as researchers continued to explore the best evaluation techniques to measure pilot performance in various aircraft (Mixon and Moroney, 1981).

One popular method is to calculate the standard deviations from the mean, such as used by Bisson et al. (1993) to estimate levels of pilot precision based on digital flight data collected from a C-141. When using standard deviations, however, information about the error relative to a given set of criteria is not provided. Instead, the standard deviation technique measures how widely values are dispersed from the average value (which may or may not be centered on the correct flight-control parameter). Root mean square errors (RMSE), on the other hand, convey the deviation from clearly established criteria and provide an assessment of how well the pilot meets various ideal standards, such as maintaining specific target altitudes, airspeeds, etc. The smaller the RMSE value, the better the performance. Various studies have used RMSE to measure pilot performance. An early study conducted by Billings, Gerke, and Wick (1975) used absolute error and RMSE variability to compare flight performance in simulated and actual flight. In another study, Caldwell and Jones (1990) used RMSE to effectively evaluate helicopter pilot performance with atropine sulfate during actual and simulated flight. As evidenced by the results of this latter study, the RMSE approach was able to accurately establish differing levels of flight-performance degradations attributable to 2-4 mg of atropine sulfate in comparison to a nominal placebo condition.

Similar to the methodology used by Caldwell and Jones (1990), this report describes a scoring routine that uses RMSE to measure the impact of fatigue (from sleep deprivation) on flight performance. This paper will discuss the methods used to objectively assess and quantify the flight performance of F-117A pilots collected using the Link F-117A Weapon System Training (WST) (L3 Communications, 1993) device at Holloman AFB. The flight methodology and examples discussed in this paper originated from a fatigue study entitled, "The Effects of 37 Hours of Continuous Wakefulness on the Alertness and Flight Performance of F-117A Pilots," in which the effects of fatigue on flight performance were objectively assessed (Caldwell et al, 2003). A subsequent, follow-on study, conducted with a drug intervention, also employed these methods to determine the extent to which a fatigue countermeasure mitigated the effects of untreated sleep loss on simulator control accuracy (Caldwell et al, 2004).

## METHODS

### Flight Simulator

The F-117A WST at Holloman Air Force Base, NM, was the simulator used to collect all of the flight performance data. The simulator is a motionless device that provides a fully-working replica of the cockpit found in the actual F-117 aircraft, including all primary and secondary flight controls, engine sounds, and cockpit lighting (L-3 Communications, 1993). It is designed to accurately replicate all aspects of the Air Force F-117A aircraft, including aircraft systems and operation. The components of the WST include the simulator itself as well as an instructor/operator station, a computer complex that includes an Alpha Server 8200 and Input/Output cabinets, and the equipment necessary for the generation of out-of-the-window and IR visual scenes. The actual F-117A aircraft is a

twin-turbofan powered, low-radar, ground-attack fighter with a single-seat cockpit. Training in the WST is directly transferable to the actual F-117 aircraft in terms of instrument flights, takeoffs and landings, instrument navigation, system operations, and air-to-ground attack procedures.

Objective flight performance data were collected using the Coherent Automated Simulation Test Environment (CoASTE) tool—a set of software routines that normally provides the capability to evaluate simulator performance, display/manipulate various data from simulator data pools, and/or trace and correct problems. The CoASTE tool's trace utility was used to capture various parameters of flight performance data at a rate of 2 Hz throughout each flight (Caldwell et al., 2003). One complete data file was generated for each simulator flight. This file contained all the data collected from the beginning to the end of the given simulation session. Each record in the file contained the time at which each data sample was collected, the actual data points themselves, and an identification field which consisted of the subject number, testing day, and testing session. The completed data files were downloaded onto a 700 MB Compact Disk Recordable (CD-R) at the conclusion of data collection before being transferred to a standard desktop Pentium-based computer where the data were further reduced and analyzed, a process that will be extensively described in this paper.

Flight Assessment

For this particular study, the primary objective was tracking the impact of fatigue on basic piloting skills. Therefore, it was crucial that a standardized flight profile be flown for all the simulator flights. The flight profile, as shown in Table 1, was comprised of 15 standardized instrument flight maneuvers that were repeated in the same sequence at 5-hour

intervals throughout the final portion of a 37-hour period of continuous wakefulness. There were 5 test flights during this time, each of which was 1-hour in duration. The flights were conducted at 2300, 0400, 0900, 1400, and 1900.

The reason the flight profile shown below was chosen was because it tapped the basic "stick and rudder" skills required of all pilots, and it permitted precise and quantifiable assessment of pilot performance across all participants in the study. If non-standardized maneuvers had been chosen, it would have been difficult to assess performance because pilots often use different, but equally correct, techniques to successfully accomplish the same mission. A pilot's ability to fly the types of precision instrument maneuvers used here enables a general assessment of how they would be able to react to more complex, real world situations.

Table 1. Flight maneuvers

| Number | Detailed maneuver descriptions |
|---|---|
| 1. | Right 360° turn at an altitude of 11,000 ft mean sea level (MSL) |
| 2. | Straight and level on a heading of 345° at 11,000 ft MSL |
| 3. | Left 360° turn at an altitude of 11,000 ft MSL |
| 4. | Straight climb from 11,000 to 13,000 ft MSL |
| 5. | Straight and level on a heading of 345° at 13,000 ft MSL |
| 6. | Descending right 360° turn to an altitude of 10,000 ft MSL |
| 7. | Straight and level on a heading of 345° |
| 8. | Left-climbing 540° turn to an altitude of 15,000 ft. |
| 9. | Straight and level on a heading of 165° at 15,000 ft MSL |
| 10. | Right 360° turn at an altitude of 15,000 ft MSL |
| 11. | Straight and level on a heading of 165° at 15,000 ft MSL |
| 12. | Left 720° turn at an altitude of 15,000 ft MSL |
| 13. | Straight descent from 15,000 to 13,000 ft MSL |
| 14. | Intercept localizer (not scored) |
| 15. | Instrument Landing System (ILS) approach to Runway 16 |

Prior to the sleep-deprivation period in which the test flights were flown, each pilot was trained to asymptote on the standardized profile via the completion of three training flights. Upon arrival for training and for participation in the study, the participant was given a pre-flight mission briefing by a member of the research staff who served as the console operator during all of the initial training flights. The research staff member informed the participant about the flight maneuvers that would be flown and the manner in which each flight would be conducted. The pilot was told that a staff member would explicitly sequence them through each of 15 flight maneuvers at the appropriate point in each of the flight profiles. The participants were instructed that they would be required to establish flight parameters within 3° of the target heading, 3 knots of the target airspeed, and 10 ft of the target altitude before each maneuver began. This was done to ensure accurate data measurement and also to increase the validity of using the performance metric of RMSE to measure flight performance. RMSE has limitations in measuring deviations from an assigned flight path when large deviations from the target value occur; therefore, it was critical that the specified parameters were met before the maneuver commenced (Hubbard, 1987). The pilot was encouraged to precisely fly all of the flight parameters and to maintain strict standards of performance throughout each of the flights.

Once the training flights were completed, the pilot participant was released and given a night off prior to the sleep-deprivation period which began on the following day. Upon return to the simulator facility (the next day), the pilot completed the 5 flights described earlier (at 2300, 0400, 0900, 1400 and 1900). This same procedure was used for all 10 pilots who were tested in the study.

DATA COLLECTION

The flight data were collected with the CoASTE tool and collected as a continuous stream from beginning to the end of each flight. The CoASTE tool captured several parameters of flight performance data (see Table 2) at a rate of 2 samples recorded every second during each flight maneuver.

Table 2. Measured simulator flight parameters

| Parameter |
|---|
| 1. Indicated altitude (feet) |
| 2. Indicated airspeed (knots) |
| 3. Indicated vertical speed (feet per min.) |
| 4. Magnetic heading (degrees) |
| 5. Pitch angle (degrees) |
| 6. Roll angle (degrees) |
| 7. Slip ((ball widths) |
| 8. Localizer/course deviation (dots) |
| 9. Glideslope/course deviation (dots) |

The data from the CoASTE tool were saved as an ASCII file and the data were then imported into Microsoft Excel for subsequent data reduction. The next step was to organize the data by splitting it into the 15 maneuvers based on the mission elapsed time that was marked by the console operator at the start of each of the 15 maneuvers throughout the flight. Once the flight data had been split into the 15 maneuvers, the next step was to trim the flight data based on the parameters for each maneuver so that the transition into and out of the maneuver was removed. For example, Maneuver 4 consisted of a straight climb at a rate of 1000 feet per minute (fpm). Some pilots began a consistent climb quickly, while others took longer to consistently climb at 1000 fpm. To eliminate any variability in transitioning from level flight into the climb, the data were trimmed so that Maneuver 4

started when the pilots displayed a consistent climb rate for several consecutive seconds and ended prior to the transition back to level flight. For Maneuver 1, a right 360° turn at a 30° angle of bank, the flight data were trimmed so that the data to be scored for the maneuver was limited to what was recorded once the pilots were consistently maintaining a 30° angle of bank. The data points leading up to the 30° mark were eliminated to reduce variability amongst pilots and to ensure that only the heart of the maneuver, and not the transition into or out of the maneuver, was scored. To eliminate the ending transition (when the pilot was rolling out of the turn) the roll-out data were excluded as well. Essentially this same type of procedure was used on all of the scored flight maneuvers.

For each maneuver, the 3 most relevant parameters were scored. This issue is further discussed below using the left 720° turn (Maneuver number 12) as an example.

Figure 1 provides a graphical representation of the raw, trimmed flight data collected on the 10 pilots for the left 720° turn across the 3 parameters that were scored for that maneuver: altitude, airspeed, and roll. On each graph, there is a single line for each individual pilot. If the pilots had flown this maneuver perfectly, there would be one solid horizontal line extending (from left to right) from the target values of 15,000 ft, 300 knots, and 30° angle of bank since every collected sample would have corresponded perfectly to the ideal values; however, varying degrees of deviations occurred amongst the pilots. The particularly large amplitude increases from the target values, present during session 3 and 4, illustrate how flight performance deteriorated over time.

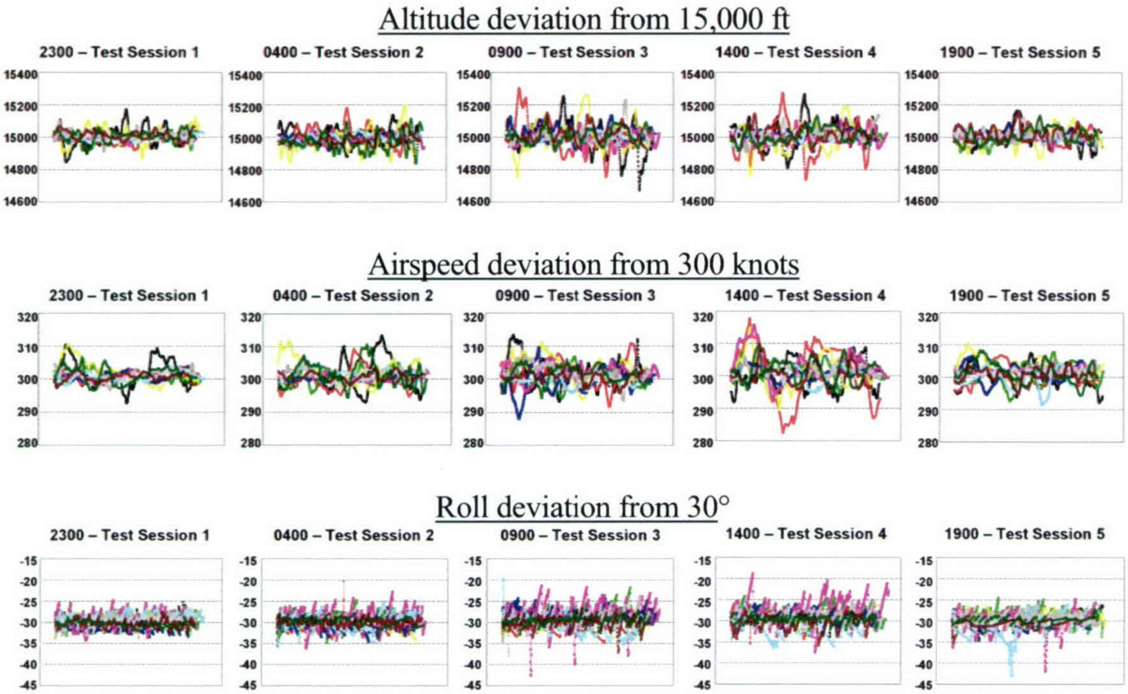



Figure 1. Ten pilots' raw data on left 720° turn maneuver across altitude, airspeed, and roll.

## DATA REDUCTION

### Root Mean Square Error (RMSE)

The next step was to calculate the RMSE for each scored flight parameter for each maneuver (the parameters are shown in Table 3). Different parameters were evaluated for each maneuver because of the relevance of those parameters to the proper conduct of specific flight tasks. For example, the altitude on a climb was constantly changing because this is the nature of a climb; therefore, altitude deviation would have been a meaningless measure in this case whereas a better parameter to score pilot control technique was the vertical velocity because it revealed a pilot's ability to maintain a consistent *climb rate*.

Table 3. Maneuver's flight parameters

| Detailed maneuver descriptions | Evaluated parameters |
|---|---|
| 1. Right 360° turn | altitude, airspeed, roll |
| 2. Straight and level 1 | altitude, airspeed, heading, roll |
| 3. Left 360° turn | altitude, airspeed, roll |
| 4. Straight climb | airspeed, vertical velocity, heading |
| 5. Straight and level 2 | altitude, airspeed, heading, roll |
| 6. Descending right 360° turn | airspeed, vertical velocity, roll |
| 7. Straight and level 3 | altitude, airspeed, heading, roll |
| 8. Left-climbing 540° turn | airspeed, vertical velocity, roll |
| 9. Straight and level 4 | altitude, airspeed, heading, roll |
| 10. Right 360° turn | altitude, airspeed, roll |
| 11. Straight and level 5 | altitude, airspeed, heading, roll |
| 12. Left 720° turn | altitude, airspeed, roll |
| 13. Straight descent | airspeed, vertical velocity, heading |
| 14. Intercept localizer | not scored |
| 15. Instrument Landing System (ILS) | |

RMSE was calculated using the following formula:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\text{Target Value} - \text{Observed Value})^2}{n}}$$

Table 4 shows how the RMSE formula was applied to a pilot's data on the left 720° turn maneuver. The observed altitude values were subtracted from the target altitude of 15,000 ft. The score was squared and summed across the entire maneuver to yield a value of 3,120,806.31. Next, the result was divided by the total number of samples (829) collected from a single pilot during the maneuver, and the square root of the result was

taken to yield the RMSE of 61.36 ft. Therefore, on this particular maneuver, the pilot averaged an error of 61.36 ft as he attempted to maintain an altitude of 15,000 ft.

Table 4. Left 720° turn – Altitude RMSE

| Time (s) | Altitude (ft) | Target Altitude (ft) | Dev ^2 |
|---|---|---|---|
| 1948.0 | 15002.9 | 15000 | 8.41 |
| 1948.5 | 15000.9 | 15000 | 0.81 |
| 1949.0 | 14999.8 | 15000 | 0.04 |
| ↓ | ↓ | ↓ | ↓ |
| 2361.0 | 14975.7 | 15000 | 590.49 |
| 2361.5 | 14973.1 | 15000 | 723.61 |
| 2362.0 | 14970.3 | 15000 | 882.09 |
| | | | |
| N | | | Sum Dev^2 |
| 829 | | | 3120806.31 |
| | | | RMSE |
| | | | 61.36 |

Figure 2 represents the RMSE values that were calculated using the raw, trimmed flight data shown in Figure 1. Thus each data line in Figure 2 has now been reduced to a single number. The RMSE values for the left 720° turn were calculated independently for altitude, airspeed, and roll, across the five testing sessions (yielding one data point per subject per session). For all the subsequent graphs (one each for the 3 scored parameters) the 10 thin lines represent the 10 pilots, while the thick line represents the composite average performance across the 10 pilots. The graphs provide a useful visual aid in viewing how errors increased over time as the pilots attempted to maintain target values of altitude, airspeed, and roll on the left 720° turn despite growing amounts of sleep deprivation. Note that this same procedure was used to characterize performance on the other maneuvers as well.

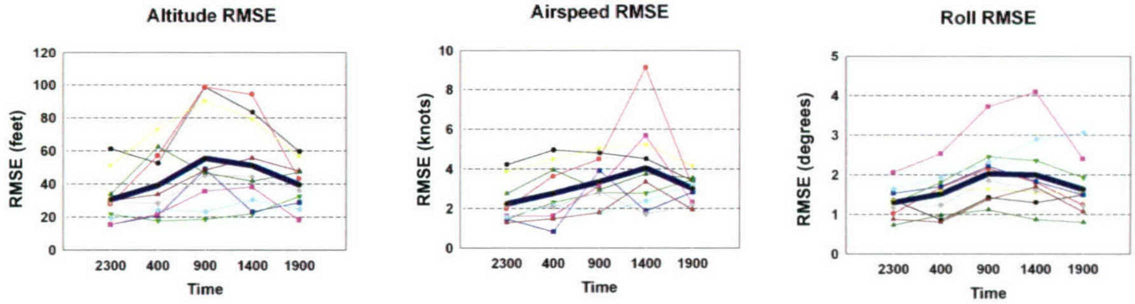



Figure 2. Ten pilots' RMSE scores on left 720° turn maneuver across altitude, airspeed, and roll.

Percent Change

Since the flight parameters were all represented in various units of measurement such as feet, knots, degrees, etc., the conversion to percent change scores enabled multiple flight parameters to be compared on the same scale. Consequently, this enabled a composite score to be calculated for each maneuver and this made it possible to compare performance across different times and across different maneuvers.

The conversion from RMSE into percent-change-from-baseline scores was accomplished using the following formula:

*Percent Change = ((Baseline RMSE – Deprivation RMSE)/Baseline RMSE) x 100.00*

For example, Table 5 shows how the altitude percent change was calculated for one pilot on the left 720° turn. The baseline scores were obtained from the third training flight conducted the day prior to the testing (sleep deprivation) period. For the 2300 session, the testing altitude RMSE of 61.36 ft was subtracted from the baseline altitude RMSE of 69.30 ft. The result was then divided by the baseline score and multiplied by 100 to yield a percent change of -11.46%. The same series of calculations were performed for the subsequent test sessions at 0400, 0900, 1400, and 1700, and this procedure was followed for all of the evaluated parameters for each maneuver.

Table 5. Left 720° turn – Altitude percent change

| Time | Baseline Altitude RMSE | Testing Altitude RMSE | Percent Change |
|---|---|---|---|
| 2300 | 69.30 | 61.36 | -11.46% |
| 400 | 69.30 | 80.77 | 16.56% |
| 900 | 69.30 | 98.66 | 42.37% |
| 1400 | 69.30 | 83.58 | 20.61% |
| 1900 | 69.30 | 59.73 | -13.81% |

The conversion of the RMSE into percent change scores revealed the decrease (or increase) in flight performance over the sleep deprivation period. Figure 3 shows the percent change scores that were calculated based on the altitude, airspeed, and roll RMSE data shown in Figure 2 (again, using the Left 720° turn as an example).

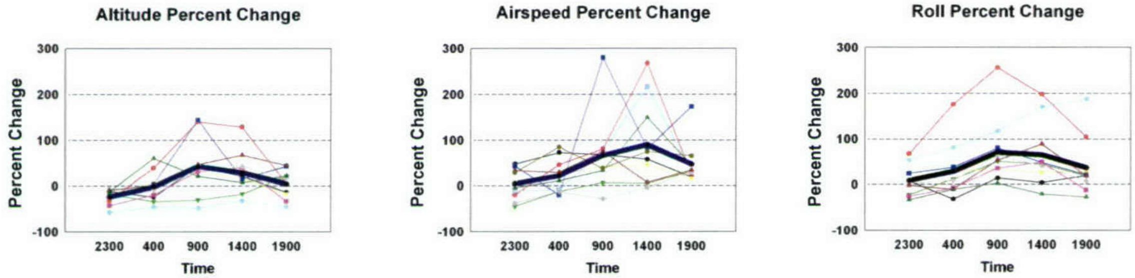



Figure 3. Ten pilots' percent change scores on left 720° turn maneuver across altitude, airspeed, and roll.

Once each scored parameter was converted from RMSE to percent change, a composite flight performance score was calculated for each of the flight maneuvers. For the left 720° turn the altitude, airspeed, and roll percent change scores were averaged together at each time point. This was accomplished for each subject. As shown in Figure 4, the composite percent change graph reveals the similar trend in performance seen in the individualized altitude, airspeed, and roll percent change graphs (Figure 3), except that greater errors are now represented by a downward deflection (all of the percent change data were inverted to aid in data interpretation (since this graphs poor performance as "down" and good performance as "up")).

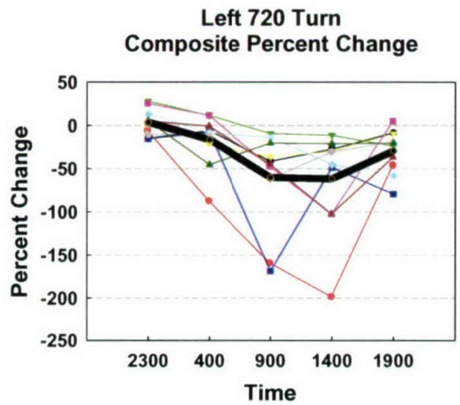
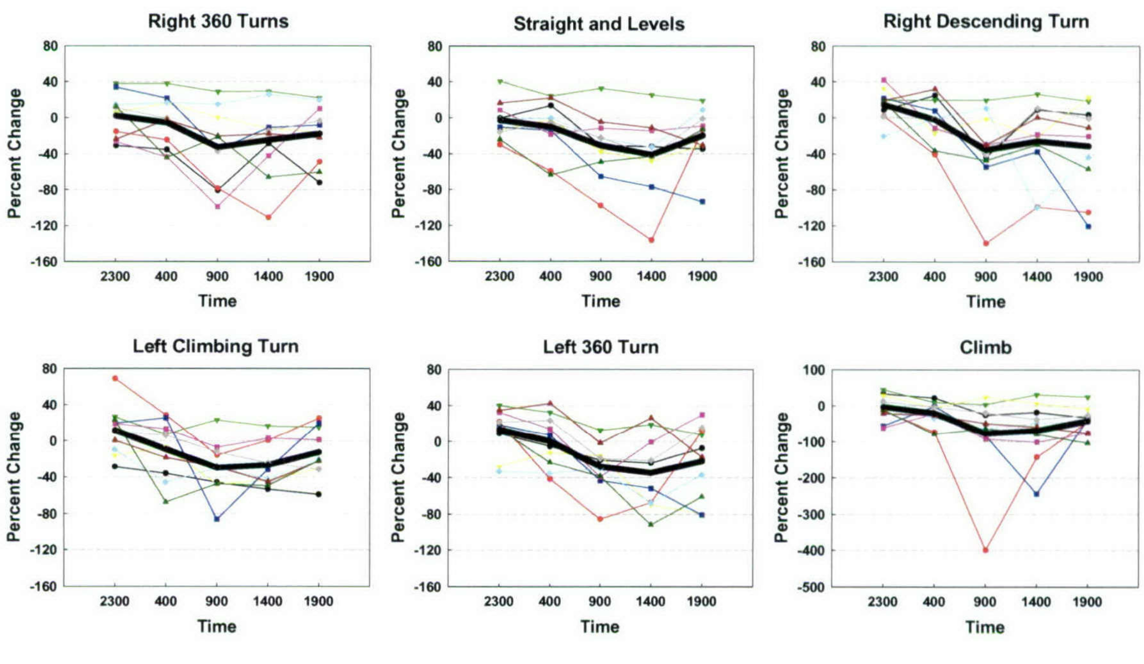



Figure 4. Ten pilots' composite percent change scores on the left 720° turn maneuver.

In addition to the left 720° turn, the composite percent change scores were calculated across the additional flight maneuvers seen in Figure 5: right 360° turns, straight and levels, right descending turn, left climbing turn, left 360° turn, climb, and descent. The same series of steps described above, with the left 720° turn, were applied to these maneuvers. Note that multiple iterations of the straight and levels and the right 360° turns have been averaged together to produce only one graph per maneuver type.



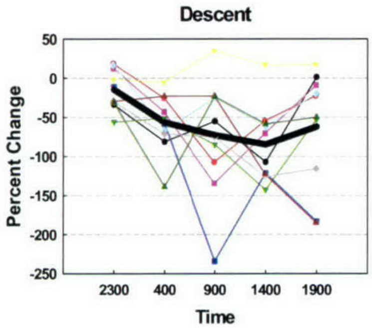



Figure 5. Ten pilots' composite percent change scores on the right 360° turns, straight and levels, right descending turn, left climbing turn, left 360° turn, climb, and descent.

To further reduce the amount of flight data, composite scores for each of the 10 pilots were calculated by averaging their individual percent change scores across the 8 maneuver types at each time point. Figure 6 depicts the 10 pilots' flight performance scores averaged across all maneuvers and reveals the individual variations in fatigue vulnerability across the pool of 10 pilots. Note that, once again, the thin lines represent individuals while the thick line represents the composite average performance across the 10 pilots.

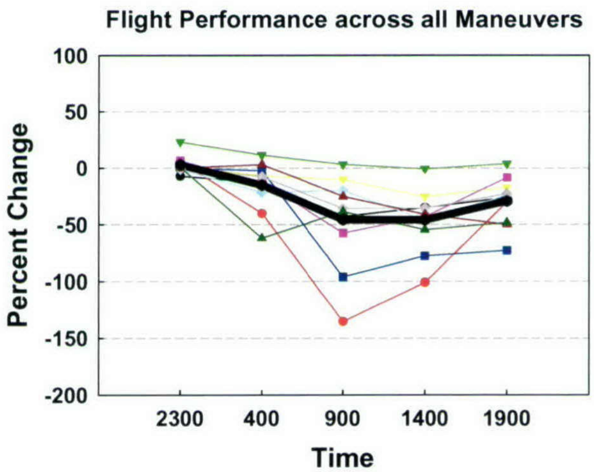



Figure 6. Ten pilots' composite percent change scores across all maneuvers.

DATA ANALYSIS

Once the flight data were transformed into percent change scores for each pilot at each time point within each maneuver, the data were analyzed with BMDP4V, Repeated Measures Analysis of Variance (ANOVA). Also, BMDP4V was used to conduct regression evaluations on all significant effects for the presence of linear, quadratic, and cubic trends. Overall, there were five testing sessions and eight different maneuver types (combined right 360° turns, combined straight and levels, right descending turn, left climbing turn, left 360° turn, climb, left 720° turn, and descent). One-way ANOVA's were performed on the eight maneuver types. Figure 7 graphically depicts the composite flight performance across the 10 pilots (averaged together) for each of the eight maneuver types, with linear, quadratic, and cubic trends annotated (bottom right corner of each graph).

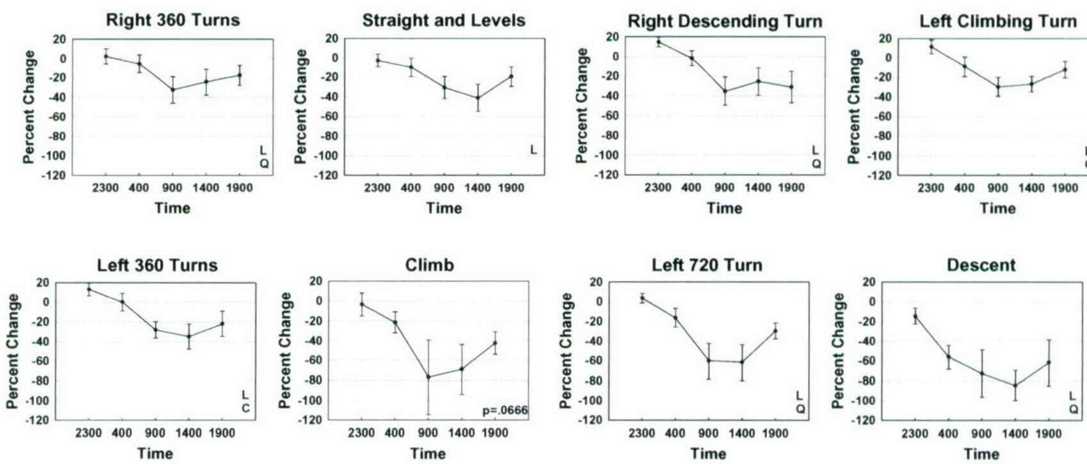



Figure 7. Ten pilots' composite percent change scores on the right 360° turns, straight and levels, right descending turn, left climbing turn, left 360° turn, climb, left 720° turn, and descent with linear, quadratic, and cubic trends annotated.

Next, an overall analysis was conducted using a two-way ANOVA in which the factors of interest were testing time and flight maneuver. As shown in Figure 8, there

was a time main effect (F(2.44,21.93)=10.72, p=.0003). Subsequent analyses indicated this to be a function of significant linear, quadratic, and cubic ($p<.05$) trends. Figure 8 reflects the general pattern seen with the various flight maneuvers, with performance accuracy seriously deteriorating after 27-32 hrs of wakefulness.

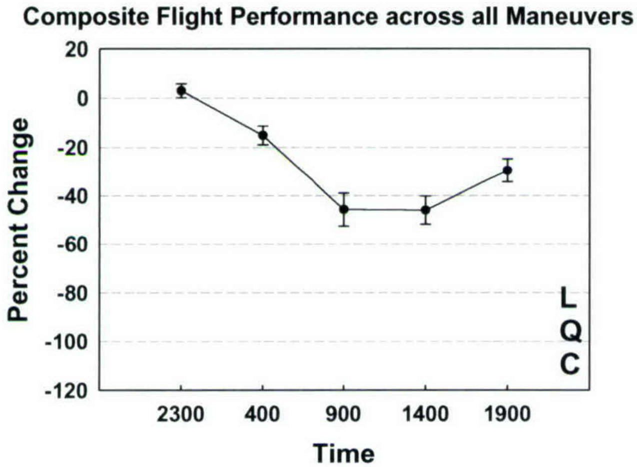



Figure 8. Ten pilots' composite flight performance across all maneuvers with linear, quadratic, and cubic trends annotated.

## CONCLUSION

As shown in this paper, the evaluation of objectively-measured flight performance requires a precisely structured methodology in order to obtain meaningful and useful results. Past researchers have developed various techniques to try to evaluate pilot performance, including using subjective ratings (Rehmann, 1982), standard deviations (Bisson et al., 1993), absolute error, and RMSE variability (Billings et al., 1975). In the case of the study outlined here, the RMSE technique was used because it permitted a clear determination of the extent of pilot performance degradations associated with lengthy periods of continuous wakefulness (37 hours). Such data are useful for substantiating the fact that this particular stressor (sleep deprivation) does in fact compromise the operational readiness of even well-trained military professionals. In the

future, these data may form the basis for comparing the effectiveness of a recommended sleep-deprivation countermeasure, and/or they may be used to help design work schedules in which critical tasks are avoided during particularly dangerous times.

Regardless of specific future applications, the procedures used to quantify these results were shown to be of significant value, and it is hoped that the descriptive procedural information provided here will enable others to take advantage of a useful analytical approach. The methodology outlined in this paper can be applied not only to the study of pilot fatigue, but to the assessment of other stressors or interventions that may either degrade or enhance aviator performance, such as chemical defense antidotes (Caldwell and Jones, 1990), alcohol (Billings et al, 1991), modafinil(Caldwell et al, 2003), or others. Clearly the computerized RMSE approach offers the advantages of objectivity, repeatability, and interpretability, while avoiding the subjectivity, unreliability, and inherent biases that can adversely impact human ratings.

REFERENCES

Billings, C. E., Demosthenes, T., White, T. R., & O'Hara, D. B. (1991). Effects of alcohol on pilot performance in simulated flight. *Aviation, Space, and Environmental Medicine, 62(3),* 233-235.

Billings, C. E., Eggspuehler, J. J., Gerke, R. J., & Chase, R. C. (1968). Studies of pilot performance: II. Evaluation of performance during low altitude flight in helicopters. *Aerospace Medicine, 39(1)*, 19-31.

Billings, C. E., Gerke, R. J., & Wick, R. L. (1975). Comparisons of pilot performance in simulated and actual flight. *Clinical Medicine, 46(3),* 304-308.

Bisson, R. U., Neville, K. J., Boll, P. A., French, J., Ercoline, W. R., McDaniel, R. L., & Storm, W. F. (1993). Digital flight data as a measure of pilot performance associated with fatigue from continuous operations during the Persian Gulf conflict.

Caldwell, J. A., Caldwell, J. L., Brown, D. L., Smythe, N. K., Smith, J. K., Mylar, J. T., Mandichak, M. L., & Schroeder, C. (2003). *The effects of 37 hours of continuous wakefulness on the physiological arousal, cognitive performance, self-reported mood, and simulator flight performance of F-117A pilots.* (USAF Research Laboratory Technical Report No. 2003-0086). Brooks City-Base, TX: Air Force Research Laboratory.

Caldwell, J. A. & H. D. Jones (1990). A comparison of computerized measurement of helicopter pilot performance with atropine sulfate during actual and simulated flight. In *AGARD Conference Proceedings: Vol 490. Safety Network to Detect Performance Degradation and Pilot Incapacitation* (pp.5-1-5-10). Neuilly Sur Seine, France: North Atlantic Treaty Organization.

Caldwell J. A., Caldwell, J. L., Smith J. K., Alvarado L., Heintz T. R., Mylar J. T., & Brown, D. L. (2004). *The efficacy of modafinil for sustaining alertness and simulator flight performance in F-117 pilots during 37 hours of continuous wakefulness*. (USAF Research Laboratory Technical Report No. 2004-0003). Brooks City-Base, TX: Air Force Research Laboratory.

Connolly, T. J., Blackwell, B.B., & Lester, L. F. (1989). A simulator-based approach to training in aeronautical decision making. *Aviation, Space, and Environmental Medicine, 60(1),* 50-52.

Fuller, J. H., Waag, W. L., & Martin, E. L. (1980). *Advanced simulator for pilot training: Design of automated performance measurement system.* (Air Force Human Resources Laboratory No. AFHRL-TR-57). Brooks City-Base, TX: Air Force Human Resources Laboratory.

Hubbard, D. C. (1987). Inadequacy of root mean square error as a performance measure (Proceedings of the Fourth International Symposium on Aviation Psychology). Columbus, Ohio.

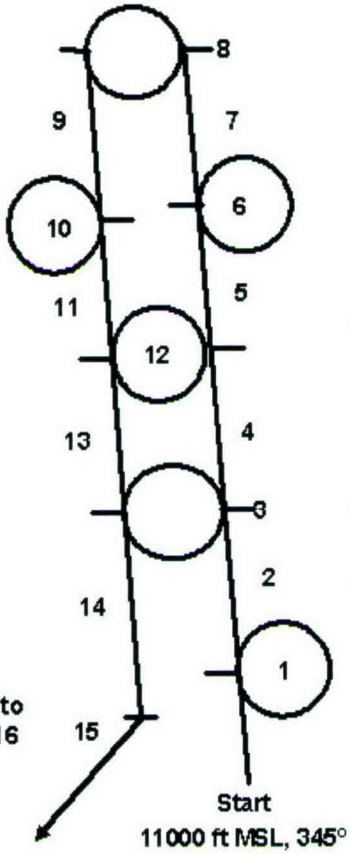
Knoop, P. A. & Welde, W. L.. (1973). *Automated pilot performance assessment in the T-37: A feasibility study* (Air Force Human Resources Laboratory AFHRL-TR-72-6). Brooks City-Base, TX: Air Force Human Resources Laboratory.

L-3 Communications. (1993). *Operation and maintenance instructions, F-117 Weapon System Trainer (WST)* (Lockheed Martin Skunk Works Technical Manual No. TO43D3-4-117-1-1). Binghamton, NY: L-3 Communications.

Mixon, T. R. & Moroney, W. F. (1981). *An annotated bibliography of objective pilot performance measures* (Naval Training Equipment Center Technical Report No. NAVTRAEQUIPCEN IH-330). Orlando, FL: Naval Training Equipment Center.

Rehmann, J. T. (1982). Pilot performance measurement: An annotated bibliography, Federal Aviation Administration Technical Center. (U.S. Department of Transportation No. DOT/FAA/EM-81/16). Washington, DC: U.S. Department of Transportation.

APPENDIX A
Flight Profile




1. 30° AOB, right 360° turn.
2. Straight and level. 2 min.
3. 30° AOB, right 360° turn.
4. Climb @ 1000 fpm to13000 MSL, 300 kts.
5. Straight and level. 2 min.
6. 30° AOB, right 360° turn descend @ 1000 fpm to10000 MSL, 300 kts.
7. Straight and level. 2 min.
8. 30° AOB, left 540° turn, climb @ 1000 fpm to 15000 MSL, 300 kts.
9. Straight and level. 2 min.
10. 30° AOB, right 360° turn.
11. Straight and level. 2 min.
12. 30° AOB, left 720° turn.
13. Descend @ 1000 fpm to 13000 MSL, 300 kts.
14. Straight and level. 2 min.
15. Turn right to 213°, descend to 6900 MSL, to intercept Rwy 16 localizer.

APPENDIX B

Holloman Simulator Fatigue Study Flight Script

Subject ID _________ Day _______ Session _______ Time _______ Date _________ Tech Initials ________

_______Turbulence 0; MEA 0; Night Illum.; ALT 11,000; ASP 300; HDG 345; Disable Autopilot/Auto-throttle

_______ SD off, IRADS off

_______ Printer paper full

_______Simulator room lights off

_______ Simulator room door closed

_______COASTE initialized, **PILOT ALT SET 30.02**

_______ Radio check

_______ Simulator canopy closed

_______EEG Eyes-Open/Eyes-Closed

_______ Take simulator off freeze/**PRINT PAGE**

_______ Initial start time of simulator __________

Stabilize flight path, straight and level, at 11,000 feet, heading 345 degrees, airspeed 300 knots.

Task 1. On my mark, your first maneuver will be a 30-degree-angle-of-bank, right 360-degree turn, from a heading of 345 degrees to 345 degrees. Maintain 11,000 feet and 300 knots…Ready…Mark. (After mark repeat: *turning right to a heading of 345 degrees, 30-degree angle of bank, at 11,000 feet, 300 knots*).

Task 2. In about 20 seconds and on my mark, your next maneuver will be 2 minutes of straight-and-level flight. Maintain a heading of 345 degrees, altitude of 11,000 feet, and airspeed of 300 knots…Ready…Mark. (After mark repeat: *maintain heading of 345 degrees, 11,000 feet, 300 knots*).

Task 3. In about 20 seconds and on my mark, your next maneuver will be a 30-degree-angle-of-bank, left 360-degree turn, from a heading of 345 degrees to 345 degrees. Maintain 11,000 feet and 300 knots… Ready…Mark. (After mark repeat: *turning left to a heading of 345 degrees, 30-degree angle of bank, at 11,000 feet, 300 knots*).

Task 4. In about 20 seconds and on my mark, your next maneuver will be a straight climb on a heading of 345 degrees from 11,000 feet to 13,000 feet at 1000 feet per minute. Maintain 300 knots…Ready…Mark. (After mark repeat: *climbing from 11,000 to 13,000 feet, 1000 feet per minute, heading 345 degrees, airspeed 300 knots*).

Task 5. In about 20 seconds and on my mark, your next maneuver will be a 2-minute straight and level. Maintain a heading of 345 degrees, altitude of 13,000 feet, and airspeed of 300 knots…Ready…Mark. (After mark repeat: *straight and level, on a heading of 345 degrees, at 13,000 feet, 300 knots*).

Task 6. In about 20 seconds and on my mark, your next maneuver will be a 30-degree-angle-of-bank, right descending turn, 360 degrees, from a heading 345 to heading 345. Descend from 13,000 feet to 10,000 feet at 1000 feet per minute. Maintain 300 knots…Ready…Mark. (After mark repeat: *30-degree-angle-of-bank, right descending 360-degree turn to a heading of 345 degrees. Descending to 10,000 feet at 1000 feet per minute, maintain 300 knots*).

Task 7. In about 20 seconds and on my mark, your next maneuver will be a 2-minute straight and level. Maintain a heading of 345 degrees, altitude of 10,000 feet, and airspeed of 300 knots…Ready…Mark. (After mark repeat: *straight and level, on a heading of 345 degrees, at 10,000 feet, 300 knots*).
(Check EEG trace)

Task 8. In about 20 seconds and on my mark, your next maneuver will be a 30-degree-angle-of-bank, left climbing turn, 540 degrees, from a heading 345 to heading 165 degrees. Climb from 10,000 feet to 15,000 feet at 1000 feet per minute. Maintain 300 knots…Ready…Mark. (After mark repeat: *30-degree-angle-of-bank, left climbing 540-degree turn to a heading of 165 degrees. Climbing to 15,000 feet at 1000 feet per minute, maintain 300 knots*).

Task 9. In about 20 seconds and on my mark, your next maneuver will be a 2-minute straight and level. Maintain a heading of 165 degrees, altitude of 15,000 feet, and airspeed of 300 knots…Ready…Mark. (After mark repeat: *straight and level, on a heading of 165 degrees, at 15,000 feet, 300 knots*).

Task 10. In about 20 seconds and on my mark, your next maneuver will be a 30-degree-angle-of-bank, right 360-degree turn, from a heading of 165 to 165 degrees. Maintain 15,000 feet and 300 knots…Ready…Mark. (After mark repeat: *turning right to a heading of 165 degrees, 30-degree angle of bank, at 15,000 feet, 300 knots*).

Task 11. In about 20 seconds and on my mark, your next maneuver will be a 2-minute straight and level. Maintain a heading of 165 degrees, altitude of 15,000 feet, and airspeed of 300 knots…Ready…Mark. (After mark repeat: *straight and level, on a heading of 165 degrees, at 15,000 feet, 300 knots*).

Task 12. In about 20 seconds and on my mark, your next maneuver will be a 30-degree-angle-of-bank, left 720-degree turn, from a heading of 165 to 165 degrees. Maintain 15,000 feet and 300 knots…Ready…Mark. (After mark repeat: *720-degree left turn to a heading of 165 degrees, 30-degree angle of bank, at 15,000 feet, 300 knots*).

Task 13. In about 20 seconds and on my mark, your next maneuver will be a straight descent on a heading of 165 degrees from 15,000 feet to 13,000 feet at 1000 feet per minute. Maintain 300 knots…Ready…Mark. (After mark repeat: *descending from 15,000 to 13,000 feet, 1000 feet per minute, maintain heading of 165 degrees, airspeed 300 knots*).

Task 14. In about 20 seconds and on my mark, your next maneuver will be a straight and level, maintaining 13,000 feet, slowing to 250 knots….Ready…Mark. (After mark repeat: *straight and level, maintaining 13,000 feet, slowing to 250 knots*).

<u>Approaching 250 knots.</u>

You are now cleared direct Higgy (destination 14) for the High ILS, Yankee, Runway 16, Low Approach, maintain 13,000 feet.

<u>Once turned towards Higgy (within 5-10 miles RNG).</u>

You are now cleared for the High ILS, Runway 16, report departing Higgy and 13,000 feet, maintaining 250 knots.

<u>Once pilot reports departing Higgy (turning to heading 213, descending out of 13,000 feet).</u>

Say: Cleared, High ILS, Runway 16.

Task 15. <u>Once the pilot intercepts the localizer.</u> **MARK**.

Report 5 miles with the gear and slow to final approach airspeed once you are configured with the gear down.

<u>Once the pilot gives the gear check.</u>

You are cleared low approach, Runway 16, report going missed approach at Decision Height. *(DH is 4283 feet)*

<u>Once the pilot reports going missed approach.</u>

**MARK** (*This is the end of the flight; give it a few seconds then put the sim on freeze and end this segment of the test by telling the pilot to go ahead and get out or standby while you get the leads disconnected.*

APPENDIX C
Holloman Fatigue Study
Participant Pre-flight Briefing

1. There will be a total of 8 flights in the study.
2. Three training flights on day 1; Five test flights on day 2.
3. You may be coached by an IP on the training flights, but not the test flights.
4. The flight profile consists of 15 instrument maneuvers including an ILS to a missed approach.
5. All maneuvers will be flown at night with no illumination.
6. Start and Stop points for each maneuver will be specified by the console operator throughout the study, so *please don't start any maneuver until instructed to do so.*
7. We will inform you about your next maneuver about 20 seconds prior to the end of the maneuver you are presently flying. However, you should continue with the current maneuver until it is complete. *You are responsible for rolling out on the correct headings and altitudes.*
8. Flight tolerances are much closer than 60-2 standards. Performance will be measured to exacting standards by computer. For example—heading within 1 degree, altitude within 1 foot, etc.
9. Autopilot and autothrottle are not to be used; however, you can use the DEP and INS for steer modes, instrument reference, and destination points.
10. The objective is to be as perfect as possible at all times.
11. Because of the EEG and Eye Tracking, you can't chew gum or candy during the flights, and you should try to keep your jaw muscles and face muscles relaxed.
12. Also, because of the EEG and Eye Tracking, conversation must be kept to a minimum at all times. It is not necessary to repeat back instructions for the maneuvers. However, during the ILS, we will ask you to call out some report points.
13. During training flights, you can talk to ask questions, get clarifications, etc. because we want you to get up-trained as well as possible on the flight profile.
14. During the testing flights, we ask that you remain as quiet as possible, but don't hesitate to ask us to repeat instructions if you forget something.
15. Make sure you have the approach book with you into the simulator. For the ILS, switch to destination 14, go to Higgy.
16. Once in the sim, we will hook up the equipment, turn off the lights, close the canopy, do a radio check, then do an eyes-open/eyes-closed EEG, take the sim off freeze, and we'll go.

Show the participant the flight profile and give him examples of the instructions.

*If the participant asks what call sign to use during the profile, tell him SIM*